NASA Contractor Report 187567
ICASE Report No. 91-43

ICASE

SINC-GALERKIN ESTIMATION OF DIFFUSIVITY IN PARABOLIC PROBLEMS

Ralph C. Smith
Kenneth L. Bowers

Contract No. NAS1-18605
May 1991

Institute for Computer Applications in Science and Engineering
NASA Langley Research Center
Hampton, Virginia 23665-5225

Operated by the Universities Space Research Association

National Aeronautics and
Space Administration

Langley Research Center
Hampton, Virginia 23665-5225

SINC-GALERKIN ESTIMATION OF DIFFUSIVITY IN PARABOLIC PROBLEMS

Ralph C. Smith[1]
Institute for Computer Applications in Science and Engineering
NASA Langley Research Center
Hampton, VA 23665

Kenneth L. Bowers
Department of Mathematical Sciences
Montana State University
Bozeman, MT 59717

ABSTRACT

A fully Sinc-Galerkin method for the numerical recovery of spatially varying diffusion coefficients in linear parabolic partial differential equations is presented. Because the parameter recovery problems are inherently ill-posed, an output error criterion in conjunction with Tikhonov regularization is used to formulate them as infinite-dimensional minimization problems. The forward problems are discretized with a sinc basis in both the spatial and temporal domains thus yielding an approximate solution which displays an exponential convergence rate and is valid on the infinite time interval. The minimization problems are then solved via a quasi-Newton/trust region algorithm. The L-curve technique for determining an appropriate value of the regularization parameter is briefly discussed, and numerical examples are given which demonstrate the applicability of the method both for problems with noise-free data as well as for those whose data contains white noise.

[1]This research was supported by the National Aeronautics and Space Administration under NASA Contract No. NAS1-18605 while the author was in residence at the Institute for Computer Applications in Science and Engineering (ICASE), NASA Langley Research Center, Hampton, VA 23665.

1 Introduction

In this paper, a fully Sinc-Galerkin method is introduced for the numerical recovery of variable diffusion coefficients in linear parabolic partial differential equations. To illustrate the method, consider the problem of estimating the spatially varying parameter $p(x)$ in the diffusion equation

$$\begin{aligned} \mathcal{L}(p)u \equiv & \frac{\partial u}{\partial t} - \frac{\partial}{\partial x}\left(p(x)\frac{\partial u}{\partial x}\right) = f(x,t), \quad 0 < x < 1 \quad t > 0 \\ & u(0,t) = u(1,t) = 0, \quad t > 0 \\ & u(x,0) = 0, \quad 0 \le x \le 1 \end{aligned} \tag{1.1}$$

given measurements of the data at the points $\{(x_p, t_q)\}_{p=1,\cdots,n_p}^{q=1,\cdots,n_q}$ in $(0,1) \times I\!R^+$. As noted in [1], problems of this type arise in applications ranging from physiological modeling to sea sediment analysis.

For many applications, it is physically reasonable to assume that p is continuous on $[0,1]$ and to let the admissible parameter set Q be defined by

$$Q = \{p \in H^1(0,1) : p(x) \ge p_0 > 0\}.$$

With this definition, the existence of a unique solution u to the forward problem can be obtained on any fixed time interval, $(0,\tau], \tau > 0$, for f sufficiently smooth.

The objective of the parameter recovery problem is to choose $p^* \in Q$ so that the solution u^* of (1.1) corresponding to p^* agrees with the true state $\hat{u}$. In general however, the true state $\hat{u}$ is not known and measurements are taken instead from an observation space Z. In this paper, the data are taken to be point evaluations and the observation space Z is defined to be $Z = I\!R^{n_p \cdot n_q}$. The observation operator $\mathcal{C} : C((0,1) \times (0,\tau]) \to Z$ is then given by

$$\mathcal{C}\psi = \{\psi(x_p, t_q)\}_{p=1,\cdots,n_p}^{q=1,\cdots,n_q}. \tag{1.2}$$

The "idealized" recovery problem may then be formulated as follows: determine $p \in Q$ so that

$$\mathcal{C}u(\cdot,\cdot,p) = \vec{d}$$

where $\vec{d}$ is used to denote the data. Since the forward problem is well-posed, the parameter recovery problem may be formulated as

$$\mathcal{K}(p) = \vec{d} \tag{1.3}$$

where the nonlinear operator $\mathcal{K}$ is defined by

$$\mathcal{K}(p) = \mathcal{C}\mathcal{L}^{-1}(p)f.$$

The problem (1.3) is impractical to solve for several reasons. As indicated in [12], the problem is ill-posed in the sense that solutions p (provided they exist) may not depend continuously on the data $\vec{d}$. Hence, discretized versions of this problem are likely to be highly ill-conditioned. Consequently, some sort of regularization (i.e., stabilization) is required to obtain an accurate approximation for p.

The regularization technique that is used is Tikhonov regularization [19] and the problem (1.3) is replaced by the minimization problem

$$\min_{p \in Q} \mathcal{T}_\alpha(p) \tag{1.4}$$

where the Tikhonov functional is

$$\mathcal{T}_\alpha(p) \equiv \frac{1}{2}\{\|\mathcal{K}(p) - \vec{d}\,\|^2 + \alpha \mathcal{J}(p)\}.$$

Here $\alpha > 0$ is a regularization parameter, which controls the tradeoff between goodness of fit to the data and stability. The penalty functional $\mathcal{J}(p)$ provides stability and allows the inclusion of *a priori* information about the true parameter p^*. Since the parameter is assumed to be "smooth" in the sense that $p \in H^1(0,1)$, the penalty functional is taken to be the norm

$$\mathcal{J}(p) = \|p\|_Q^2 \equiv \int_0^1 [p'(x)]^2 v(x)dx + \epsilon \int_0^1 [p(x)]^2 v(x)dx. \tag{1.5}$$

The parameter ϵ is of the order 10^{-6} and the weight v is taken to be the positive function $v(x) = x(1-x)$. The reasons for weighting the integral as well as including the second term and enforcing $\mathcal{J}(p)$ to be strictly positive will be discussed in the fourth section of this paper. By using arguments similar to those in [8] and [15] and assuming that $\mathcal{K}(p)$ is one to one, it can be shown that with this definition for $\mathcal{J}(p)$, the solutions p_α to (1.4) converge as

the regularization parameter $\alpha \to 0$ and as the perturbations in the data and operator tend to zero.

Due to the infinite dimensionality of Q and that of the state space, the problem (1.4) is an infinite-dimensional minimization problem. In order to develop a practical numerical scheme, the problem must be replaced by a sequence of finite-dimensional problems; that is, one must approximate the operator $\mathcal{K}$ and minimize the functional T_α over a finite-dimensional admissible subspace of Q.

The evaluation of $\mathcal{K}(p)$ requires the solution of the partial differential equation (1.1). Similar PDE's must be solved to obtain the components of the derivative $\mathcal{K}'(p)$. The construction of an approximate solution to these forward problems commonly begins with a Galerkin discretization of the spatial variable with time-dependent coefficients. This yields a system of ordinary differential equations which is solved via differencing techniques. Due to stability constraints on the discrete evolution operator, low-order methods with small time steps are often required to obtain accurate approximations. Moreover, this time-stepping must be repeated at each step in the minimization of (1.4). A final difficulty lies in the need to interpolate at data points which do not coincide with the nodes of the ODE solver.

In contrast, the method of this work implements a Galerkin scheme in time as well as space thus bypassing many of the difficulties associated with time-stepping methods in the context of inverse problems. This possibility was first explored in [12]. In contrast to the methods of that work however, both the spatial and temporal basis functions are taken to be compositions of sinc functions with suitable conformal maps.

By discretizing the forward problem in this manner, the optimal exponential convergence rate is exhibited throughout the infinite time domain, even in the presence of boundary singularities. The validity and exponential convergence rate of the approximate solution throughout all time is especially important in those problems in which the data is sampled at large temporal values t_q. Furthermore, the sinc quadrature rules yield coefficient matrices which are efficiently constructed for the forward problem and easily updated when the forward techniques are employed in a parameter recovery scheme. The efficiency of the inverse scheme is further augmented by the fact that the component matrices used in formulating the finite-dimensional penalty functional are identical to those used when constructing the

forward coefficient matrices and hence need to be formed only once. The efficiency and accuracy of the forward solver and the ease of formulating the penalty functional are then manifested in the inverse algorithm for a large class of problems.

The foundations of the Sinc-Galerkin method and the fundamental quadrature rules are described in Section 2. A thorough review of sinc function properties can be found in [17] and [18]. In the third section of this paper, the Sinc-Galerkin system for the forward problem is constructed and implementation details are discussed. The section closes with the discussion of a very robust and accurate algorithm for solving the resulting algebraic system. Section 4 includes the finite-dimensional minimization problem with the discussion centering around the construction of the various components of the Tikhonov functional. By taking advantage of sinc function properties, efficient routines for approximating the nonlinear operator $\mathcal{K}(p)$ and the penalty functional $\mathcal{J}(p)$ are developed. In the next section, a quasi-Newton/trust region scheme is outlined for solving the finite-dimensional minimization problem. The paper concludes with a section containing numerical examples. Of the many examples tested, those discussed in this section best exhibit the features necessary for the practical implementation of the Sinc-Galerkin method. A brief discussion of the Generalized Cross Validation (GCV) and L-curve techniques for choosing the regularization parameter α is given at the beginning of the section, and the applicability of these techniques in conjunction with the Sinc-Galerkin method is demonstrated by the numerical results. Finally, results are included both from data sets with white noise and from sets to which no noise was added. As shown in these examples, the Sinc-Galerkin method works equally well in both cases.

2 Sinc Function Properties

For the Sinc-Galerkin method, the basis functions are derived from the Whittaker cardinal (sinc) function

$$\text{sinc}(x) \equiv \frac{\sin(\pi x)}{\pi x}, \quad -\infty < x < \infty$$

and its translates

$$S(k,h)(x) \equiv \text{sinc}\left(\frac{x-kh}{h}\right), \quad h > 0.$$

For $h^* = \frac{\pi}{4}$, three adjacent members of this sinc family $(S(k,h^*)(x), k = -1, 0, 1)$ are shown in Figure 1.

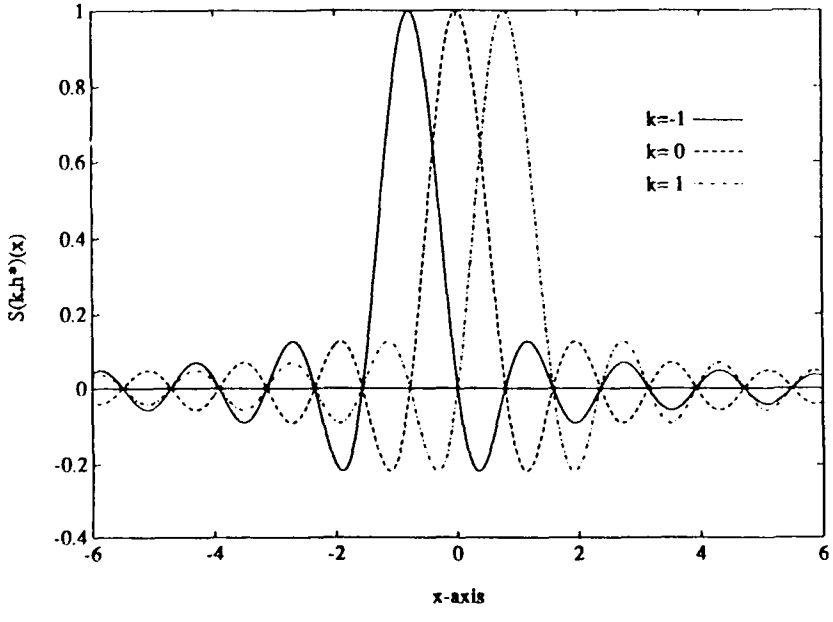

Figure 1. Three Adjacent Members $\left(S(k,h^*)(x), k = -1, 0, 1, h^* = \frac{\pi}{4}\right)$ of the Translated Sinc Family.

To construct basis functions on the intervals (0,1) and $(0,\infty)$, respectively, consider the conformal maps

$$\phi(z) = ln\left(\frac{z}{1-z}\right) \tag{2.1}$$

and

$$\Upsilon(w) = ln(w). \tag{2.2}$$

The map ϕ carries the eye-shaped region

$$D_E = \left\{ z = x + iy : \left| arg\left(\frac{z}{1-z}\right)\right| < d \leq \frac{\pi}{2} \right\} \tag{2.3}$$

onto the infinite strip

$$D_S = \{\xi = \zeta + i\eta : |\eta| < d \leq \frac{\pi}{2}\}. \tag{2.4}$$

Similarly, the map Υ carries the infinite wedge

$$D_W = \{w = t + is : |arg(w)| < d \leq \frac{\pi}{2}\} \tag{2.5}$$

onto the strip D_S. These regions are depicted in Figure 2.

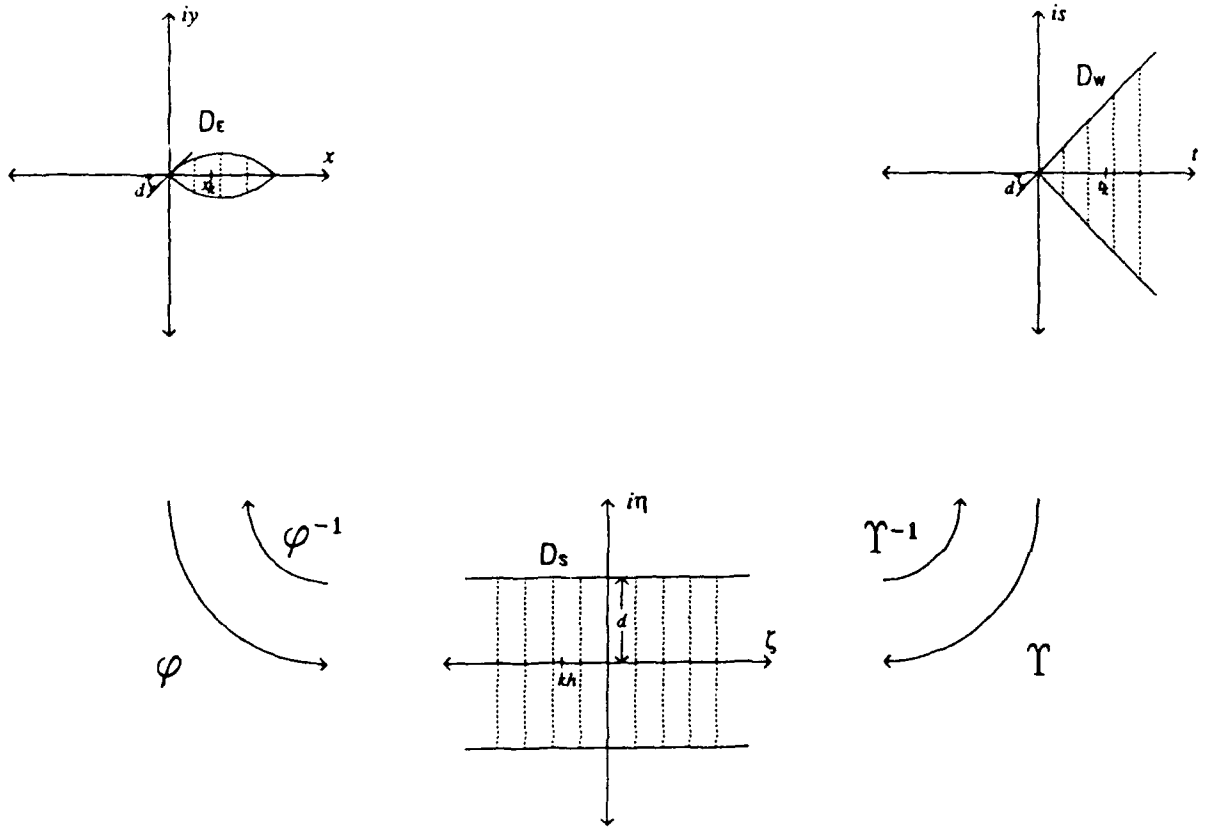

Figure 2. The Domains D_S, D_E, and D_W.

The sinc gridpoints $z_k \in (0,1)$ in D_E will be denoted x_k since they are real. Similarly, the gridpoints $w_k \in (0,\infty)$ in D_W will be denoted t_k. Both are inverse images of the equispaced grid in D_S; that is,

$$x_k = \phi^{-1}(kh) = \frac{e^{kh}}{1+e^{kh}} \tag{2.6}$$

and

$$t_k = \Upsilon^{-1}(kh) = e^{kh}. \tag{2.7}$$

To simplify notation throughout the remainder of this section, the pairs ϕ, D_E and Υ, D_W are referred to generically as χ, D. It is understood that the subsequent definition, theorems, and identities hold in either setting. Furthermore, the inverse of χ is denoted by ψ.

The important class of functions for sinc interpolation and quadrature is denoted $B(D)$ and defined next.

Definition 2.1. *Let $B(D)$ be the class of functions F which are analytic in D, satisfy*

$$\int_{\psi(t+L)} |F(z)dz| \to 0, \quad t \to \pm\infty$$

where $L = \{is : |s| < d \leq \frac{\pi}{2}\}$, and on the boundary of D (denoted ∂D) satisfy

$$N(F) \equiv \int_{\partial D} |F(z)dz| < \infty.$$

The following theorem for functions in $B(D)$ is found in [16].

Theorem 2.1. *Let Γ be $(0,1)$ or $(0,\infty)$ when $\chi = \phi$ or Υ, respectively. If $F \in B(D)$ and $z_j = \psi(jh) = \chi^{-1}(jh)$, $j = 0, \pm 1, \pm 2, \cdots$, then for $h > 0$ sufficiently small*

$$\left| \int_\Gamma F(z)dz - h \sum_{j=-\infty}^{\infty} \frac{F(z_j)}{\chi'(z_j)} \right| \leq K_1 e^{-2\pi d/h}. \tag{2.8}$$

Theorem 2.1 illustrates the exponential convergence rate which is a trademark of the sinc methods. There is a common occasion when it is possible to evaluate the infinite series appearing in (2.8), namely when integrating against $S(k,h) \circ \chi$. In general, however, the series must be truncated. With additional hypotheses, it is proven in [11] and [17] that the truncation need not be at the expense of the exponential convergence.

Theorem 2.2. *Assume $F \in B(D)$ and that there exist positive constants K, α, and β such that*

$$\left|\frac{F(\tau)}{\chi'(\tau)}\right| \leq K \begin{cases} e^{-\alpha|\chi(\tau)|}, & \tau \in \psi((-\infty, 0)) \\ e^{-\beta|\chi(\tau)|}, & \tau \in \psi([0, \infty)). \end{cases} \tag{2.9}$$

Then for h sufficiently small

$$\left|\int_\Gamma F(z)dz - h \sum_{j=-M}^{N} \frac{F(z_j)}{\chi'(z_j)}\right| \leq K_1 e^{-2\pi d/h} + \frac{K}{\alpha} e^{-\alpha M h} + \frac{K}{\beta} e^{-\beta N h}. \tag{2.10}$$

Theorems 2.1 and 2.2 are used to establish a uniform error bound when constructing an approximate solution to the forward second-order time-dependent problems. The application of these quadrature theorems is facilitated by the identities

$$\delta_{pi}^{(0)} \equiv [S(p,h) \circ \chi(z)]\Big|_{z=z_i} = \begin{cases} 1, & i = p \\ 0, & i \neq p, \end{cases} \tag{2.11}$$

$$\delta_{pi}^{(1)} \equiv h\left[\frac{d}{d\chi} S(p,h) \circ \chi(z)\right]\Big|_{z=z_i} = \begin{cases} 0, & i = p \\ \dfrac{(-1)^{i-p}}{(i-p)}, & i \neq p \end{cases} \tag{2.12}$$

and

$$\delta_{pi}^{(2)} \equiv h^2\left[\frac{d^2}{d\chi^2} S(p,h) \circ \chi(z)\right]\Big|_{z=z_i} = \begin{cases} -\dfrac{\pi^2}{3}, & i = p \\ \dfrac{(-2)(-1)^{i-p}}{(i-p)^2}, & i \neq p \end{cases} \tag{2.13}$$

which denote the evaluation at the gridpoint z_i of the sinc-map compositions and their derivatives with respect to the map χ.

3 The Forward Problem

Consider the second-order parabolic problem

$$\begin{aligned} \mathcal{L}u(x,t) \equiv \frac{\partial u}{\partial t} - \frac{\partial}{\partial x}\left(p(x)\frac{\partial u}{\partial x}\right) &= f(x,t), \quad 0 < x < 1\ ,\ t > 0 \\ u(0,t) = u(1,t) &= 0, \quad t > 0 \\ u(x,0) &= 0, \quad 0 \le x \le 1\ . \end{aligned} \tag{3.1}$$

To define the Sinc-Galerkin approximation to (3.1), let $S_i(x) \equiv S(i,h_x) \circ \phi(x)$ and $S_j^*(t) \equiv S(j,h_t) \circ \Upsilon(t)$, and take the basis to be $\{S_{ij}\}_{i=-M_x,\cdots,N_x}^{j=-M_t,\cdots,N_t}$ where

$$S_{ij}(x,t) \equiv S_i(x)S_j^*(t).$$

The approximate solution is then defined by way of the tensor product expansion

$$u_{m_x m_t}(x,t) = \sum_{i=-M_x}^{N_x}\sum_{j=-M_t}^{N_t} u_{ij}S_{ij}(x,t), \qquad \begin{aligned} m_x &= M_x + N_x + 1 \\ m_t &= M_t + N_t + 1\ . \end{aligned} \tag{3.2}$$

The $m_x \cdot m_t$ unknown coefficients $\{u_{ij}\}$ are determined by orthogonalizing the residual with respect to the set of sinc functions $\{S_p S_q^*\}_{p=-M_x,\cdots,N_x}^{q=-M_t,\cdots,N_t}$. This yields the discrete Galerkin system

$$(\mathcal{L}u_{m_x m_t} - f, S_p S_q^*) = 0 \tag{3.3}$$

for $p = -M_x,\cdots,N_x$ and $q = -M_t,\cdots,N_t$. The inner product $(\cdot,\cdot)$ is taken to be

$$(F,G) = \int_0^\infty \int_0^1 F(x,t)G(x,t)w(x,t)dxdt \tag{3.4}$$

with the weight

$$w(x,t) = w(x)w^*(t) = (\phi'(x))^{-\frac{1}{2}}(\dot{\Upsilon}(t))^{\frac{1}{2}}\ . \tag{3.5}$$

A thorough discussion motivating this choice of weight can be found in [10] and [13].

Because of the tensor nature of the approximate solution, the domain on which (3.1) is posed, and the form of the inner product, the discrete system (3.3) can be formulated by combining the discrete systems corresponding to the one-dimensional problems

$$\begin{aligned} \dot{u}(t) &= r(t) \qquad , \ 0 < t < \infty \\ u(0) &= 0 \end{aligned} \tag{3.6}$$

and

$$\begin{aligned} (p(x)u'(x))' &= r(x) \qquad , \ 0 < x < 1 \\ u(0) &= u(1) = 0 \ . \end{aligned} \tag{3.7}$$

This latter approach also illustrates the sinc parameter selections which are needed when implementing the method.

Continuing with (3.6), a discrete system is formed by orthogonalizing the residual $\dot{u}_{m_t}(t) - r(t)$ with respect to $\{S_q^*\}_{q=-M_t}^{N_t}$. Before invoking the quadrature rules, integration by parts is used to transfer the differentiation of u onto $S_q^*\sqrt{\dot{\Upsilon}}$, where again, $w^* = \sqrt{\dot{\Upsilon}}$ denotes the temporal inner product weight. To guarantee that the boundary terms vanish, it is assumed that

$$\lim_{t\to 0^+} \frac{u(t)}{\sqrt{t}} = \lim_{t\to\infty} \frac{u(t)}{\sqrt{t}} = 0.$$

Finally, the resulting integrals are evaluated via (2.10) or (2.8) when possible. With respect to (2.9), the condition

$$\left| u(t)\sqrt{\dot{\Upsilon}(t)} \right| \le L \begin{cases} t^{\gamma}, & t \in (0,1) \\ t^{-\delta}, & t \in [1,\infty) \end{cases}$$

guarantees the boundedness necessary to truncate the infinite quadrature rule. With γ and δ specified and M_t chosen, the parameter selections

$$h_t = \sqrt{\frac{\pi d}{\gamma M_t}}$$

and

$$N_t = \left\lfloor \frac{\gamma}{\delta} M_t + 1 \right\rfloor \tag{3.8}$$

where $\lfloor \cdot \rfloor$ denotes the greatest integer function, balance the asymptotic quadrature errors in (2.10) to at least $\mathcal{O}(e^{-(\pi d \gamma M_t)^{\frac{1}{2}}})$.

In many parabolic systems, it is reasonable to assume that the solution decays exponentially at infinity, that is that the solution satisfies

$$\left| u(t)\sqrt{\dot{\gamma}(t)} \right| \le L \begin{cases} t^{\gamma}, & t\epsilon(0,1) \\ e^{-\delta t}, & t\epsilon[1,\infty) \end{cases}$$

or, more succinctly,

$$|u(t)| \leq Kt^{\gamma+\frac{1}{2}}e^{-\delta t}. \tag{3.9}$$

Under this supposition, Lund [11] shows that the condition (3.8) can be replaced by

$$N_t = \left[\frac{1}{h_t} ln\left(\frac{\gamma}{\delta} M_t h_t\right) + 1\right]. \tag{3.10}$$

The selection N_t in (3.10) significantly reduces the size of the discrete system with no loss of accuracy. It is also noted that the size of the discrete system and the expected error are dictated by the asymptotic behavior of u at the endpoints.

The discrete system for (3.6) can then be formulated as follows. Let $I^{(\ell)}$, $\ell = 0, 1$ denote the $m_t \times m_t$ matrices whose qj-th entry is $\delta_{qj}^{(\ell)}$ from (2.11) and (2.12) and let $\mathcal{D}(\eta)$ be the diagonal matrix with entries $\eta(t_{-M_t}), \cdots, \eta(t_{N_t})$. The vector of unknowns $\vec{u} = [u_{-M_t}, \cdots, u_{N_t}]^T$ is then related to the known vector $\vec{r} = [r(t_{-M_t}), \cdots, r(t_{N_t})]^T$ by

$$A_t \vec{u} = \mathcal{D}((\dot{\Upsilon})^{-\frac{1}{2}})\vec{r} \tag{3.11}$$

where

$$A_t = \left[-\frac{1}{h_t} I^{(1)} + \frac{1}{2} I\right] \mathcal{D}((\dot{\Upsilon})^{\frac{1}{2}}). \tag{3.12}$$

Further details concerning the derivation of the system (3.11) can be found in [10] and a thorough analysis of the spectrum of A_t is given in [3].

The preceding discussion applied to the problem (3.7) follows a similar development. The map Υ of (2.2) is replaced by the map ϕ of (2.1) (since (3.7) is posed in the interval (0,1)) and h_t is replaced by h_x. Orthogonalizing the residual and two integrations by parts yields the system

$$\int_0^1 u(x)\left[p(x)\left(S_p(x)\frac{1}{\sqrt{\phi'(x)}}\right)'\right]' dx = \int_0^1 r(x)S_p(x)\frac{1}{\sqrt{\phi'(x)}}dx \tag{3.13}$$

for $p = -M_x, \cdots, N_x$. To guarantee that the boundary terms vanish, it is assumed that

$$\left(pu'\frac{S_p}{\sqrt{\phi'}}\right)(x)\Big|_0^1 = \left(pu\left(\frac{S_p}{\sqrt{\phi'}}\right)'\right)(x)\Big|_0^1 = 0.$$

In anticipation of the parameter recovery problem which motivates this analysis, the term $p(x)$ in (3.13) is expanded as a linear combination of sinc functions with two Hermite like

algebraic terms added to accommodate the nonzero values of p at $x = 0$ and $x = 1$. The finite-dimensional approximation of p then takes the form

$$\begin{aligned} p_{m_x}(x) &= c_{-M_x}(1-x) + c_{N_x}x + \sum_{k=-M_x+1}^{N_x-1} c_k S_k(x) \\ &\equiv \sum_{k=-M_x}^{N_x} c_k \zeta_k(x) \ . \end{aligned} \tag{3.14}$$

In the forward problem, the coefficients $\{c_k\}_{k=-M_x}^{N_x}$ are known whereas in the corresponding parameter recovery problem, they are unknown and are determined via methods to be discussed in Section 4. The number of basis functions used in the expansion is chosen so as to guarantee a square coefficient matrix. This is done to simplify the implementation of the method when applied to the PDE (3.1) of interest.

The expansion (3.14) is substituted into (3.13) and the resulting integrals are evaluated via (2.10) or (2.8) when possible. As shown in [13], the decay condition (2.9) equates to the condition

$$|u(x)P(x)| \le L \begin{cases} x^{\alpha+\frac{1}{2}}, & x \in (0, \frac{1}{2}) \\ (1-x)^{\beta+\frac{1}{2}}, & x \in [\frac{1}{2}, 1) \end{cases}$$

where

$$P(x) \equiv p(x) - p(0)(1-x) - p(1)x.$$

This may be replaced by the more stringent condition

$$|u(x)P(x)| \le K x^{\alpha+\frac{1}{2}}(1-x)^{\beta+\frac{1}{2}}. \tag{3.15}$$

The asymptotic errors are then balanced by the choices

$$h_x = \sqrt{\frac{\pi d}{\alpha M_x}}$$

and

$$N_x = \left[\!\left[\frac{\alpha}{\beta} M_x + 1 \right]\!\right]$$

where $[\![\cdot]\!]$ again denotes the greatest integer function. Note that if $\frac{\alpha}{\beta}M_x$ is an integer, this integer can be selected for N_x.

With $\vec{u}, \vec{r}$, and $I^{(\ell)}$, $\ell = 0, 1, 2$, defined as before, the system for (3.6) may be written as

$$A(p)\vec{u} = \mathcal{D}((\phi')^{-\frac{3}{2}})\vec{r} \tag{3.16}$$

where

$$\begin{aligned} A(p) &= \left[-\frac{1}{h_x^2} I^{(2)} + \frac{1}{4} I^{(0)}\right] \mathcal{D}((\phi')^{\frac{1}{2}}) \mathcal{D}(\vec{p}_\psi) \\ &- \left[\frac{1}{h_x} I^{(1)} - \mathcal{D}\left(\frac{2x-1}{2}\right)\right] \mathcal{D}((\phi')^{-\frac{1}{2}}) \mathcal{D}(\vec{p}_{\psi'}) \, . \end{aligned} \tag{3.17}$$

The notation $\mathcal{D}(\vec{p}_\psi)$ and $\mathcal{D}(\vec{p}_{\psi'})$ denotes the diagonal matrices containing the components of the vectors $\vec{p}_\psi$ and $\vec{p}_{\psi'}$ which are defined as follows. First

$$\vec{p}_\psi = \Psi \vec{c}$$

where $\vec{c} = [c_{-M_x}, \cdots, c_{N_x}]^T$ and Ψ has the block structure

$$\Psi = [\vec{\psi}_L \vdots I^{(0)} \vdots \vec{\psi}_R]_{m_x \times m_x}$$

with

$$\vec{\psi}_L = [(1 - x_{-M_x}), \cdots, (1 - x_{N_x})]^T$$

and

$$\vec{\psi}_R = [x_{-M_x}, \cdots, x_{N_x}]^T.$$

Again, the $m_x \times (m_x - 2)$ matrix $I^{(0)}$ has components $\delta^{(0)}_{qj}$ as defined in (2.11) with $-M_x \leq q \leq N_x$ and $-M_x + 1 \leq j \leq N_x - 1$. Also,

$$\vec{p}_{\psi'} = \Psi' \vec{c}$$

where

$$\Psi' = [-\vec{1} \vdots \, -\frac{1}{h_x} \mathcal{D}(\phi') I^{(1)} \vdots \vec{1}]_{m_x \times m_x}.$$

Here $\vec{1} = [1, \cdots, 1]^T$, $\mathcal{D}(\phi')$ is $m_x \times m_x$, and $I^{(1)}$ is $m_x \times (m_x - 2)$ with components $\delta^{(1)}_{qj}$ as defined in (2.12).

As shown in [13], the system (3.16) yields an approximate solution which exhibits exponential convergence to the solution u of (3.7). Further details concerning the derivation of the system as well as additional quadrature hypotheses can be found in this reference.

The above results for the one-dimensional problems (3.6) and (3.7) can then be pieced together to form the Sinc-Galerkin system for the time-dependent parabolic problem (3.1). The resulting discrete system is built from the matrices A_t (in (3.12)) and $A(p)$ (in (3.17)) of the one-dimensional problems. The parameter selections are still necessary and all that remains is to asymptotically balance the resulting errors from each one-dimensional problem.

When the decay conditions (3.9) and (3.15) are combined to yield

$$|P(x)u(x,t)| \leq Kx^{\alpha+\frac{1}{2}}(1-x)^{\beta+\frac{1}{2}}t^{\gamma+\frac{1}{2}}e^{-\delta t}, \tag{3.18}$$

then the choices

$$h_x = \sqrt{\frac{\pi d}{\alpha M_x}},$$

$$h_t = h_x,$$

$$N_x = \left\llbracket \frac{\alpha}{\beta} M_x + 1 \right\rrbracket,$$

$$M_t = \left\llbracket \frac{\alpha}{\gamma} M_x + 1 \right\rrbracket,$$

and

$$N_t = \left\llbracket \frac{1}{h_t} ln\left(\frac{\gamma}{\delta} M_t h_t\right) + 1 \right\rrbracket,$$

for the stepsizes and summation limits balance the asymptotic errors. If one takes $d = \frac{\pi}{2}$, then the above choices yield an asymptotic error rate of order $\mathcal{O}(e^{-\pi\sqrt{\alpha M_x/2}})$ for the quadratures.

Given M_x, N_x, M_t, N_t, and $h \equiv h_x = h_t$ as defined above, the discrete system for (3.1) is

$$A(p)U\mathcal{D}\left((\dot{\Upsilon})^{-\frac{1}{2}}\right) + \mathcal{D}\left((\phi')^{-\frac{3}{2}}\right)UA_t^T = G \tag{3.19}$$

where

$$G = \mathcal{D}\left((\phi')^{-\frac{3}{2}}\right)F\mathcal{D}\left((\dot{\Upsilon})^{-\frac{1}{2}}\right).$$

The diagonal matrices $\mathcal{D}\left((\phi')^{-\frac{3}{2}}\right)$ and $\mathcal{D}\left((\dot{\Upsilon})^{-\frac{1}{2}}\right)$ have sizes $m_x \times m_x$ and $m_t \times m_t$, respectively. The $m_x \times m_t$ matrices U and F contain the unknowns $\{u_{ij}\}$ and the known values $f(x_i, t_j)$.

The discrete Sinc-Galerkin system (3.19) can then be solved for U via a generalized Schur decomposition (page 396 of [6]). As guaranteed by the results of Moler and Stewart [14], there exist unitary matrices Q_1, Z_1, Q_2, and Z_2 such that

$$
\begin{aligned}
Q_1^* A(p) Z_1 &= P \\
Q_1^* \mathcal{D}\left((\phi')^{-\frac{3}{2}}\right) Z_1 &= R \\
Q_2^* \mathcal{D}\left((\dot{\Upsilon})^{-\frac{1}{2}}\right) Z_2 &= S \\
Q_2^* A_t Z_2 &= T
\end{aligned}
$$

where P, R, S, and T are upper triangular. If $Y = Z_1^* U Z_2$ and $C = Q_1^* G Q_2$, then (3.19) transforms to

$$PYT^* + RYS^* = C.$$

By comparing the k-th columns, one finds that

$$P \sum_{j=k}^{n} t_{kj} \vec{y}_j + R \sum_{j=k}^{n} s_{kj} \vec{y}_j = \vec{c}_k$$

which yields

$$(t_{kk} P + s_{kk} R) \vec{y}_k = \vec{c}_k - P \sum_{j=k+1}^{n} t_{kj} \vec{y}_j - R \sum_{j=k+1}^{n} s_{kj} \vec{y}_j \tag{3.20}$$

(for convenience, it is assumed that all matrices are $n \times n$ and indexed from 1 to n). With the assumption that the matrix $(t_{kk} P + s_{kk} R)$ is nonsingular, the solution to (3.20) is easily found by recursively solving triangular systems.

Although this algorithm does require complex algebra, it is quite efficient and requires no assumptions concerning the diagonalizability of the component matrices. It should be noted that a "real" version of this algorithm also exists [5]. In this latter algorithm, Q_1, Z_1, Q_2, and Z_2 are orthogonal with P, S quasi-upper triangular and R, T upper triangular. As proven in [5], the real algorithm is extremely stable and numerical tests have indicated that the complex algorithm described above is also robust.

4 The Finite-Dimensional Minimization Problem

As noted in the introduction, the minimization problem

$$\min_{p \in Q} \mathcal{T}_\alpha(p)$$

where

$$\mathcal{T}_\alpha(p) = \frac{1}{2}\{\|\mathcal{K}(p) - \vec{d}\,\|^2 + \alpha\|p\|_Q^2\} \tag{4.1}$$

is infinite-dimensional and hence must be replaced by a sequence of finite-dimensional problems before a viable numerical scheme may be developed. Following from (3.14), the approximating admissible parameter sets are taken to be

$$Q_{m_x} = \left\{ p_{m_x} : \; p_{m_x}(x) = \sum_{k=-M_x}^{N_x} c_k \zeta_k(x) \right\}$$

where

$$\zeta_k(x) = \begin{cases} 1 - x, & k = -M_x \\ S_k(x), & -M_x + 1 \le k \le N_x - 1 \\ x, & k = N_x \end{cases} \tag{4.2}$$

and $S_k(x) \equiv S(k, h_x) \circ \phi(x)$. The associated finite-dimensional optimization problem can then be formulated as

$$\min_{p_{m_x} \in Q_{m_x}} \hat{T}_\alpha(p_{m_x}) \tag{4.3}$$

for

$$\hat{T}_\alpha(p_{m_x}) \equiv \frac{1}{2}\left\{\|\hat{K}(p_{m_x}) - \vec{d}\,\|^2 + \alpha\|p_{m_x}\|_Q^2\right\}. \tag{4.4}$$

The approximation $\hat{K}(p_{m_x}) : \; \mathbb{R}^{m_x} \to \mathbb{R}^{n_p \cdot n_q}$ to $\mathcal{K}(p)$ is obtained by applying the point evaluation operator C in (1.2) to $u_{m_x m_t}$ in (3.2). If the set of observation points $\{(x_p, t_q)\}_{p=1,\cdots,n_p}^{q=1,\cdots,n_q}$ can be represented as a tensor product of spatial and temporal points, then $\hat{K}(p_{m_x})$ has the representation

$$\hat{K}(p_{m_x}) = C\ \overrightarrow{co(U)} \tag{4.5}$$

where the matrix U solves (3.19). The matrix concatenation $\overrightarrow{co(U)}$ is the vector in $\mathbb{R}^{m_x \cdot m_t}$ which is obtained by successively stacking the columns of the $m_x \times m_t$ matrix U. C is an

$(n_p \cdot n_q) \times (m_x \cdot m_t)$ evaluation matrix which can be formulated as follows. Define the $n_p \times m_x$ spatial evaluation matrix E_x to have components

$$[E_x]_{p,i} = S_i(x_p), \quad 1 \le p \le n_p, \quad -M_x \le i \le N_x$$

and define the $n_q \times m_t$ temporal evaluation matrix E_t to have components

$$[E_t]_{q,j} = S_j^*(t_q), \quad 1 \le q \le n_q, \quad -M_t \le j \le N_t.$$

Then

$$C = E_t \otimes E_x.$$

It is noted that if the set of observation points is not rectangular as described above, then point evaluation can be done directly via (3.2). This latter option is less efficient however, than that defined in (4.5).

The discrete penalty functional $\|p_{m_x}\|_Q^2$ is formed by substituting the expansion (3.14) into the definition (1.5). This yields

$$\|p_{m_x}\|_Q^2 = \int_0^1 [p'_{m_x}(x)]^2 v(x) dx + \varepsilon \int_0^1 [p_{m_x}(x)]^2 v(x) dx \approx \vec{c}^{\,T} Q \vec{c}$$

where the $m_x \times m_x$ matrix $Q = Q_d + Q_f$ has components

$$[Q_d]_{k\ell} \approx \int_0^1 \zeta'_k(x) \zeta'_\ell(x) v(x) dx, \quad -M_x \le k, \ell \le N_x$$

and

$$[Q_f]_{k\ell} \approx \varepsilon \int_0^1 \zeta_k(x) \zeta_\ell(x) v(x) dx, \quad -M_x \le k, \ell \le N_x.$$

The matrix entries are approximations in the sense that sinc quadrature is used to evaluate many of the integrals.

For the choice of basis functions in (4.2), the matrix Q_d is given by

$$Q_d = \begin{bmatrix} \frac{1}{6} & \vec{q}_d{}^{\,T} & -\frac{1}{6} \\ \vec{q}_d & \hat{Q}_d & -\vec{q}_d \\ -\frac{1}{6} & -\vec{q}_d{}^{\,T} & \frac{1}{6} \end{bmatrix}.$$

Integration by parts and the application of the sinc quadrature formula (2.10) yields the $(m_x - 2) \times (m_x - 2)$ matrix

$$\hat{Q}_d = -\frac{1}{h_x} I^{(2)}$$

where again $I^{(2)}$ denotes the matrix whose qj-th entry is $\delta_{qj}^{(2)}$ from (2.13). The zeroing of all other quadrature terms is a result of the choice $v(x) = \frac{1}{\phi'(x)} = x(1-x)$. The $(m_x - 2) \times 1$ vector $\vec{q}_d$ has components

$$[\vec{q}_d]_k = h_x(x_k - 3x_k^2 + 2x_k^3), \quad -M_x + 1 \leq k \leq N_x - 1$$

and is again obtained via sinc quadrature. Because $I^{(2)}$ is negative definite (see [16]), the matrix Q_d is nonnegative definite. The zero eigenvalue results from the fact that the first and last columns of Q_d differ only in sign.

Direct integration and sinc quadrature are also used to obtain the matrix

$$Q_f = \varepsilon \begin{bmatrix} \frac{1}{20} & \vec{q}_{fl}{}^T & -\frac{1}{30} \\ \vec{q}_{fl} & \hat{Q}_f & \vec{q}_{fr} \\ -\frac{1}{30} & \vec{q}_{fr}{}^T & \frac{1}{20} \end{bmatrix}.$$

Here

$$\hat{Q}_f = h_x \mathcal{D}(x^2(1-x)^2)$$

where $\mathcal{D}(\eta)$ again denotes the diagonal matrix with entries $\eta(x_{-M_x}), \cdots, \eta(x_{N_x})$. The vectors $\vec{q}_{fl}$ and $\vec{q}_{fr}$ have components

$$[\vec{q}_{fl}]_k = h_x x_k^2 (1 - x_k)^3$$

and

$$[\vec{q}_{fr}]_k = h_x x_k^3 (1 - x_k)^2$$

for $k = -M_x + 1, \cdots, N_x - 1$. The matrix Q_f is strictly positive definite.

Although the matrix Q is full, it is very efficient to construct since the Toeplitz matrix $I^{(2)}$ is also needed in the forward solver. For $\varepsilon > 0$, Q is symmetric and positive definite and hence has a Cholesky decomposition $Q = R^T R$ where R is upper triangular. It then follows that the penalty term $\|p_{m_x}\|_Q^2$ yields the quadratic form

$$\vec{c}^{\,T} R^T R \vec{c} = \|R\vec{c}\,\|^2 \tag{4.6}$$

where $\|\cdot\|$ denotes the Euclidean norm. As will be shown in the next section, this factorization admits a particularly useful diagonalization of the corresponding finite-dimensional minimization problem. It also facilitates the plotting of the L-curve to determine a suitable regularization parameter α (see Section 6 and [7]).

5 The Trust Region Scheme

In the discussion of this section, it is useful to highlight the dependence of the operators in (4.4) on the unknown vector $\vec{c} = [c_{-M_x}, \cdots, c_{N_x}]^T$ (see (3.14)). Letting

$$K(\vec{c}\,) = \hat{K}(p_{m_x}(\vec{c}\,)) = C\,\overrightarrow{co(U)}$$

and noting (4.6), the optimization problem (4.3) may be replaced by

$$\min_{\vec{c} \in R^{m_x}} T_\alpha(\vec{c}\,) \tag{5.1}$$

where

$$T_\alpha(\vec{c}\,) \equiv \frac{1}{2}\{\|K(\vec{c}\,) - \vec{d}\,\|^2 + \alpha\|R\vec{c}\,\|^2\}.$$

To obtain a minimizer for the nonlinear functional T_α, a quasi-Newton/trust region scheme is used (see [2]).

The basis for this approach is the iteration

$$\vec{c}_{k+1} = \vec{c}_k + \vec{s}_k$$

where $\vec{s}_k$ solves the constrained minimization problem

$$\min_{\vec{s}_k \in R^{m_x}} \frac{1}{2}\{\|K(\vec{c}_k) + K'(\vec{c}_k)\vec{s}_k - \vec{d}\,\|^2 + \alpha\|R(\vec{c}_k + \vec{s}_k)\|^2\} \tag{5.2}$$

subject to $\|R\vec{s}_k\| \leq \delta_k$. The trust region radius δ_k is chosen so that the quadratic model adequately reflects the behavior of T_α within the trust region; that is, δ_k is chosen so that

$$T_\alpha(\vec{c}_k + \vec{s}_k) \approx \frac{1}{2}\{\|K(\vec{c}_k) + K'(\vec{c}_k)\vec{s}_k - \vec{d}\,\|^2 + \alpha\|R(\vec{c}_k + \vec{s}_k)\|^2\}$$

whenever $\|R\vec{s}_k\| \leq \delta_k$. The minimization problem (5.2) is solved using an approach similar to that in [9]. The problem is first diagonalized using the Singular Value Decomposition (SVD). With the change of variables

$$\bar{s} = R\vec{s}_k,$$

the objective functional in (5.2) becomes

$$\frac{1}{2}\{\|A\bar{s} - \bar{b}\|^2 + \alpha\|\bar{c} + \bar{s}\|^2\}$$

where $A = K'(\vec{c}_k)R^{-1}$, $\bar{b} = \vec{d} - K(\vec{c}_k)$ and $\bar{c} = R\vec{c}_k$. Let A have the SVD

$$A = UDV^T$$

where $U_{(n_p \cdot n_q)\times(n_p \cdot n_q)}, V_{m_x \times m_x}$ are orthogonal and

$$[D_{(n_p \cdot n_q)\times m_x}]_{i,j} = \begin{cases} \sigma_i, & \text{if } i = j \text{ and } i \leq m_x \\ 0, & \text{otherwise.} \end{cases}$$

Here σ_i denotes a singular value of A. The second change of variables

$$\hat{s} = V^T\bar{s}, \quad \hat{b} = U^T\bar{b}, \quad \hat{c} = V^T\bar{c}$$

yields the diagonalized problem

$$\min_{\hat{s}\in R^{m_x}} \frac{1}{2}\{\|D\hat{s} - \hat{b}\|^2 + \alpha\|\hat{c} + \hat{s}\|^2\} \tag{5.3}$$

subject to $\|\hat{s}\| \leq \delta_k$.

The theory of constrained optimization is used to solve (5.3). By the Kuhn-Tucker criterion [4], there exists a Lagrange multiplier $\mu \geq 0$ such that

$$D^T(D\hat{s} - \hat{b}) + \alpha(\hat{c} + \hat{s}) + \mu\hat{s} = 0. \tag{5.4}$$

From (5.4) it follows that (5.3) has a unique solution of the form

$$\hat{s} = \hat{s}(\mu) = \{D^T D + (\alpha + \mu)I\}^{-1}(D^T \hat{b} - \alpha \hat{c}).$$

If $\|\hat{s}(0)\| < \delta_k$, then the constraint in (5.3) is not active and $\vec{s} = R^{-1}V\hat{s}(0)$ solves (5.2); otherwise the constraint is active and the solution to (5.2) is given by $\vec{s} = R^{-1}V\hat{s}(\mu)$ where $\mu \geq 0$ is the unique solution to

$$g(\mu) \equiv \|\hat{s}(\mu)\| - \delta_k = 0. \tag{5.5}$$

An approximate solution to the scalar equation (5.5) is then determined via the hook step algorithm in [2] (see page 134). This algorithm requires both $g(\mu)$ and $g'(\mu)$. As shown in [9], the function $g(\mu)$ can be expanded as

$$g(\mu) = \left[\sum_{j=1}^{m_s} \left(\frac{\sigma_j \hat{b}_j - \alpha \hat{c}_j}{\sigma_j^2 + \alpha + \mu}\right)^2\right]^{\frac{1}{2}} - \delta_k \tag{5.6}$$

when $\hat{b}_j$ and $\hat{c}_j$ are components of $\hat{b}$ and $\hat{c}$, respectively. The derivatives $g'(\mu)$ are easily obtained from the form (5.6).

The trust region radius δ_k in (5.3) is chosen so that $T_\alpha(\vec{c}\,)$ has sufficient decrease at each iteration so as to guarantee convergence to a local minimizer of T_α. This is accomplished via the updating algorithm in [2] (page 143) with the decay requirement taken to be

$$T_\alpha(\vec{c}_k + \vec{s}_k) \leq T_\alpha(\vec{c}_k) + \tilde{\varepsilon}\nabla T_\alpha(\vec{c}_k)^T \vec{s}_k$$

with $\tilde{\varepsilon} = 10^{-4}$.

An important numerical issue in the implementation of the trust region scheme is the formulation of the derivation of the operator K. Here the derivative, or Jacobian, is an $(n_p \cdot n_q) \times m_x$ matrix whose ν-th column is given by

$$[K'(\vec{c}\,)]_\nu = \lim_{\mathcal{T}\to 0} \frac{1}{\mathcal{T}}[K(\vec{c} + \mathcal{T}\hat{e}_\nu) - K(\vec{c}\,)]$$

where the standard unit vector $\hat{e}_\nu$ has components

$$[\hat{e}_\nu]_k = \delta_{\nu k} = \begin{cases} 1 & \text{if } k = \nu,\ -M_x \leq k \leq N_x \\ 0 & \text{otherwise.} \end{cases}$$

In the examples that are presented in Section 6, the Jacobians were calculated with a standard forward difference scheme. This scheme is easy to implement and accurate enough for the purposes of the method. If further efficiency is desired, a directional derivative scheme such as that described in [12] can be used. For this method, the trade-off for the added efficiency is an algorithm which is more difficult to implement and a slight loss of accuracy in some cases.

6 Implementation and Numerical Examples

The four examples reported in this section were selected from a large collection of problems to which the Sinc-Galerkin method was applied. The results are representative of those obtained on other sample problems.

The first example demonstrates the application of the Sinc-Galerkin method to a model problem in which the state solution was sampled directly; that is, no external noise was added to the data. To demonstrate the feasibility of the method for problems with noisy data, the same problem is revisited in Example 5.2 but with pseudo-random white noise added to the data. In Example 5.3, the parameter to be recovered has a logarithmic boundary singularity at $x = 0$ while the parameter in Example 5.4 is the shifted Gaussian function that was considered in [12]. In all four examples, the dynamics of the problem are assumed to be modeled by (1.1) with the forcing function $f(x,t)$ consistent with the state solution $u(x,t) = x(1-x)\sin(4\pi x)t^2e^{-t}$ and the true diffusion parameter p. In each case, $d = \frac{\pi}{2}$ (see (2.3) and (2.5)).

The errors for the method are reported on the set of uniform gridpoints

$$\begin{aligned} U &= \{z_0, z_1, \cdots, z_{100}\} \\ z_k &= k\ell, \quad k = 0, 1, \cdots, 100 \end{aligned} \tag{6.1}$$

with stepsize $\ell = \frac{1}{100}$. With p and p_{m_x} denoting the true and approximate parameters respectively, the errors are reported as

$$\|p_U(\ell)\| = \max_{0 \le k \le 100} |p(z_k) - p_{m_x}(z_k)|.$$

The error and convergence results are tabulated in the form $.aaa - \gamma$ which represents $.aaa \times 10^{-\gamma}$. All problems were run with sixteen place accuracy on a Vax 8550.

A very important practical consideration is the choice of the regularization parameter α for a given (error contaminated) data set. One would like to choose α so that $\|p - p_\alpha\|$ is minimized, where p_α denotes the α-dependent unknown diffusion coefficient. If the error in the data is random, then under certain conditions (see [20]) the method of Generalized Cross Validation (GCV) yields a statistical estimate of the size of $\|\mathcal{K}(p) - \mathcal{K}(p_\alpha)\|$ which is related to $\|p - p_\alpha\|$. For Tikhonov regularization, this estimate is given by the GCV functional

$$V(\alpha) = \frac{\frac{1}{n}\|K(\vec{c}_\alpha) - \vec{d}\|^2}{\left[\frac{n - m_x}{n} + \sum_{i=1}^{m_x} \frac{\alpha}{(\sigma_i^2 + n\alpha)}\right]^2} \tag{6.2}$$

where $\vec{c}_\alpha$ solves (5.1). Here $n = n_q \cdot n_p$ denotes the number of data points and $\{\sigma_i\}$ are the singular values of the operator $K'(\vec{c}_\alpha)R^{-1}$. To approximate the value of α which minimizes $\|p - p_\alpha\|$, one attempts to solve the minimization problem

$$\min_{\alpha \ge 0} V(\alpha).$$

Because the GCV method requires the singular values of $K'(\vec{c}_\alpha)$, it is relatively expensive to implement when m_x and n are large. A second disadvantage to this method for choosing the regularization parameter is that the GCV plots are often very flat making it difficult to determine a minimum value of $V(\alpha)$ and hence an optimal value of α (see Figure 6 in the next section). Finally, one often has optimization settings in which the GCV hypotheses are not satisfied.

A second method for determining the regularization parameter is to plot the norm of the penalty functional, $\|R\vec{c}_\alpha\|$, versus the norm of the residual, $\|K(\vec{c}_\alpha) - \vec{d}\,\|$ (see [7]). In this way one can qualitatively get an idea of the compromise between the minimization of these two quantities. The scheme for determining the "optimal" regularization parameter consists of finding those values of α such that $(\|K(\vec{c}_\alpha) - \vec{d}\,\|,\ \|R\vec{c}_\alpha\|)$ lies in the "corner" of the resulting curve, known as the L-curve. This method for choosing the regularization parameter α is easy to apply and often gives more conclusive results than the GCV method. Both methods are illustrated in the examples.

In all four examples, the $m_x \times 1$ initial vector $\vec{c}_0 = [.5, .5, \cdots, .5, .5]^T$ was used. Several other positive startup vectors were also tried with similar results in each case. Hence the method seems to be quite robust with respect to the choice of the initial vector.

Finally, in the examples the symbol α is used to denote both the regularization parameter (see (5.1)) and the sinc decay parameter (see (3.18)). The use of this symbol for both quantities is well established in the literature and thus difficult to avoid in this setting. It should be obvious from the context, however, which quantity is being discussed and there should be no ambiguity resulting from the dual use of this symbol.

Example 6.1 In this example, the true diffusion parameter is taken to be the analytic function $p(x) = 1 + \sin(\pi x)$. Since the state solution is $u(x,t) = x(1-x)\sin(4\pi x)t^2e^{-t}$, the decay condition (3.18) yields the choices $\alpha = \beta = 2$, $\gamma = \frac{3}{2}$, and $\delta = 1$. The data was sampled on a regular grid $\{(x_p, t_q)\} \subset (0,1) \times (0,2]$. Nine equally spaced points $x_p = p\Delta x, \Delta x = .1$, were taken in space and four equally spaced temporal points $t_q = q\Delta t$, $\Delta t = .5$, were taken for a total of $n = 54$ data points. No noise was added so the data consisted of direct measurements of the state solution. For varying values of the regularization parameter α, the L-curve is plotted in Figure 3. Note that the values $\alpha = 10^{-7}$ through $\alpha = 10^{-11}$ yield points $(\|K(\vec{c}_\alpha) - \vec{d}\,\|, \|R\vec{c}_\alpha\|)$ in the "corner" of the curve. The uniform errors for $\alpha = 10^{-8}$ are reported in Table 1 with the first four columns indicating the index limits for the expansion of the state variable and fifth column indicating the number of basis functions used in the expansion of p_{m_x}. The convergence of the method is demonstrated both by the results in the last column of Table 1 and by Figure 4 which shows the true and approximate diffusion parameters with $\alpha = 10^{-8}$.

M_x	N_x	M_t	N_t	m_x	$\|p_u(\ell)\|$
8	8	10	4	17	$.7414 - 0$
16	16	21	7	33	$7648 - 1$
24	24	39	9	49	$.3111 - 1$

Table 1. Errors on the Uniform Grid U with $\alpha = 10^{-8}$ in Example 6.1.

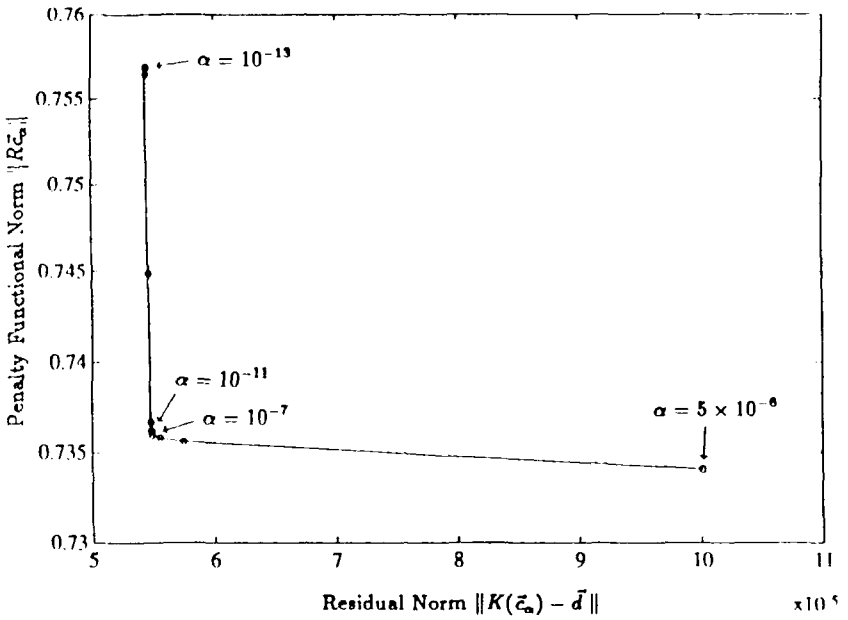

Figure 3. The Tikhonov L-curve for Example 6.1.

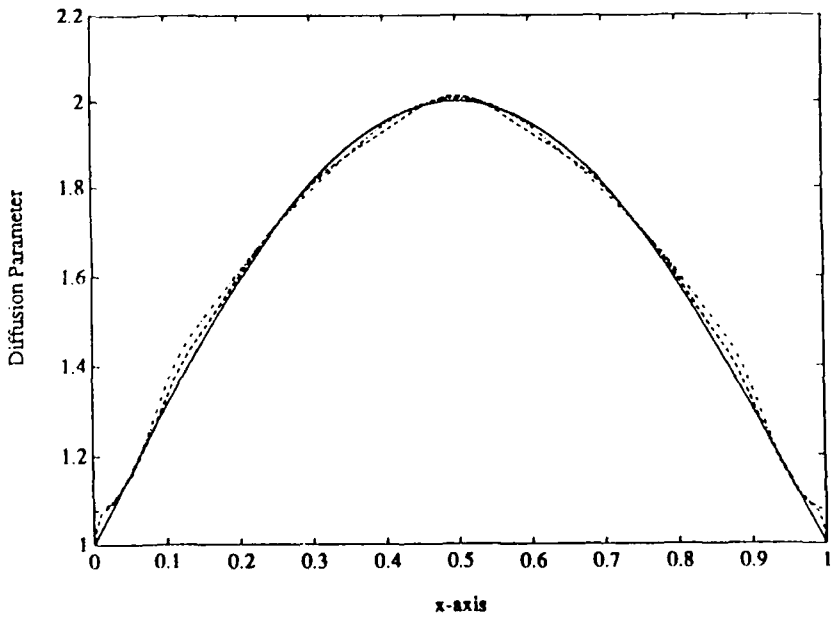

Figure 4. True and Approximate Diffusion Parameters for Example 6.1 with $\alpha = 10^{-8}$ $\cdot - \cdot$ $(M_x = 16)$, $- - -$ $(M_x = 24)$, —— (True).

Example 6.2 Here the true parameter and state solution are the same as those in Example 6.1, and hence $p(x) = 1 + \sin(\pi x)$ and $u(x,t) = x(1-x)\sin(4\pi x)t^2e^{-t}$. The same observation points were used but to this data however, we added a pseudo-random noise vector ε from a Gaussian distribution with mean 0 and standard deviation σ chosen so that the noise-to-signal ratio $\sigma/\|d\| = 0.001$ (noise - 0.1% of the signal). The L-curve and GCV curves are plotted in Figures 5 and 6, respectively. Note that the values $\alpha = 10^{-5}$ through $\alpha = 5 \times 10^{-8}$ yield points $(\|K(\vec{c}_\alpha) - \vec{d}\,\|, \|R\vec{c}_\alpha\|)$ in the "corner" of the L-curve whereas all values of α less than 10^{-5} yield apparent minima of the GCV curve. For $M_x = 16$, the uniform errors obtained with $\alpha = 10^{-3}$, $\alpha = 10^{-6}$, and $\alpha = 10^{-9}$ are reported in Table 2. Corresponding plots of the true and approximate parameters are shown in Figure 7. Note that the "corner" value $\alpha = 10^{-6}$ provides a good choice for the regularization parameter whereas $\alpha = 10^{-9}$ is not large enough to damp out the contribution due to the smaller singular values. This latter observation can be predicted from the L-curve but less easily from the GCV plot. Finally, the choice $\alpha = 10^{-3}$ causes too much smoothing and information about the parameter is lost. By comparing the results in Tables 1 and 2, it can be seen that the error in this example with $\alpha = 10^{-6}$ and $M_x = 16$ is virtually the same as the error in Example 6.1 with $\alpha = 10^{-8}$ and $M_x = 16$. The results from this example demonstrate the viability of the method for problems with noisy data.

	$\alpha = 10^{-3}$	$\alpha = 10^{-6}$	$\alpha = 10^{-9}$
$\|p_U(\ell)\|$	$.2658 - 0$	$.7737 - 1$	$.4357 - 0$

Table 2. Errors on the Uniform Grid U with $M_x = 16$ in Example 6.2.

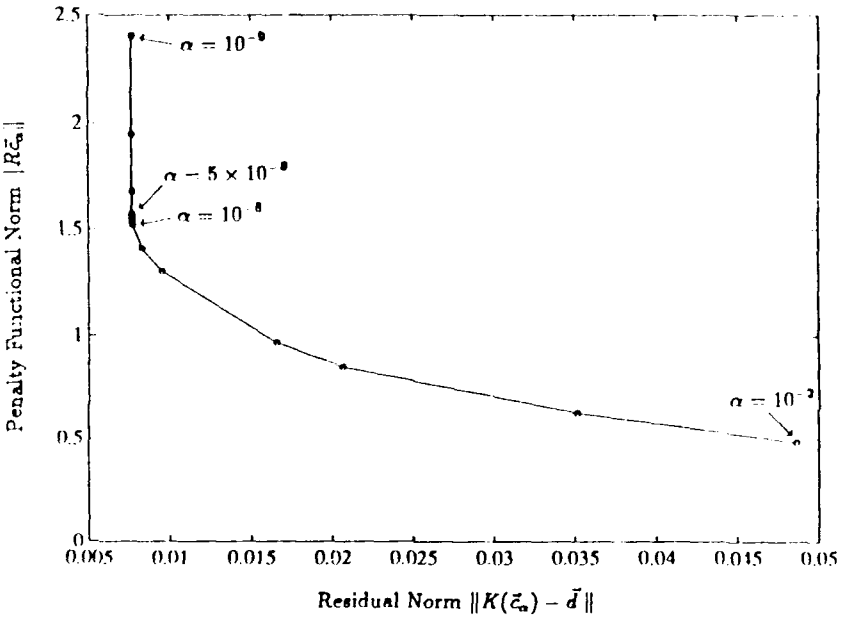

Figure 5. The Tikhonov *L*-curve for Example 6.2.

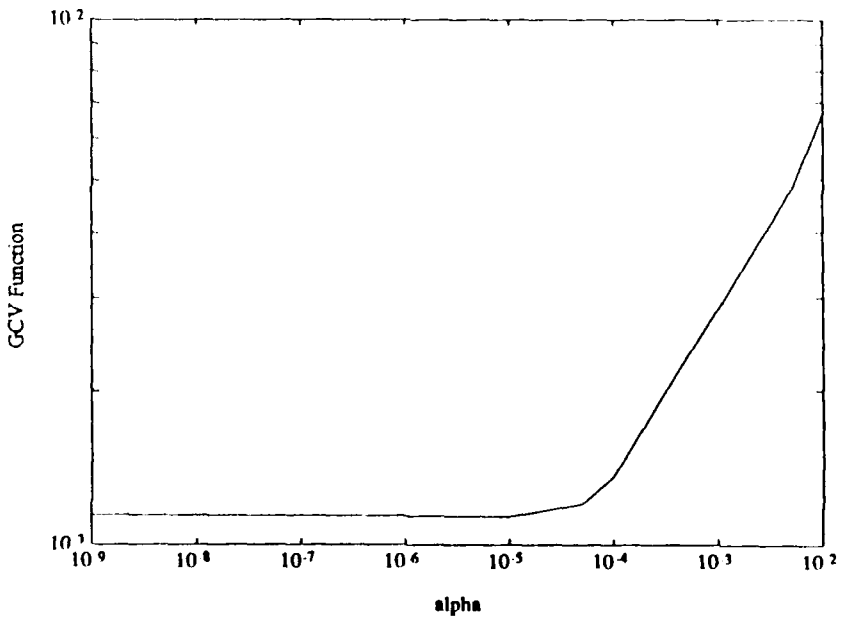

Figure 6. The GCV Functional for Example 6.2.

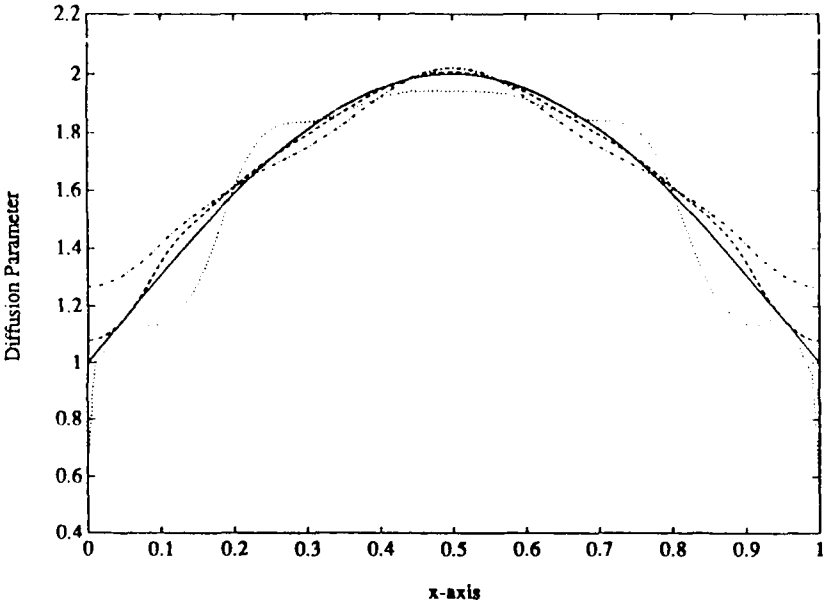

Figure 7. True and Approximate Diffusion Parameters for Example 6.2 with $M_x = 16$ $- \cdot -$ ($\alpha = 10^{-3}$), $- - -$ ($\alpha = 10^{-6}$), $\cdots$ ($\alpha = 10^{-9}$), —— (True).

Example 6.3 The true parameter in this example is $p(x) = 1 + \frac{1}{2}x + xln(x)$ which has a logarithmic singularity at $x = 0$. As before, the state solution is $u(x,t) = x(1-x)\sin(4\pi x)t^2e^{-t}$ and thus the decay parameters $\alpha = \beta = 2, \gamma = \frac{3}{2}$, and $\delta = 1$ are consistent with the condition (3.18). To demonstrate the method for another set of observation points, nineteen equally spaced points $x_p = p\Delta x$, $\Delta x = .05$, were taken in space and four equally spaced temporal points $t_q = q\Delta t, \Delta t = .5$ were taken for a total of $n = 72$ data points. No noise was added so the data consisted of direct measurements of the state solution. Since the L-curve was nearly identical to that of Example 6.1, the regularization parameter was taken to be $\alpha = 10^{-8}$. The uniform errors for this choice are reported in Table 3 and the true and approximate parameters are shown in Figure 8. Both the table and the figure demonstrate the convergence of the method in spite of the logarithmic singularity in the diffusion parameter.

M_x	N_x	M_t	N_t	m_x	$\|p_u(\ell)\|$
8	8	10	4	17	$1.2171 - 0$
16	16	21	7	33	$0.1330 - 0$
24	24	39	9	49	$0.7285 - 1$

Table 3. Errors on the Uniform Grid U with $\alpha = 10^{-8}$ in Example 6.3.

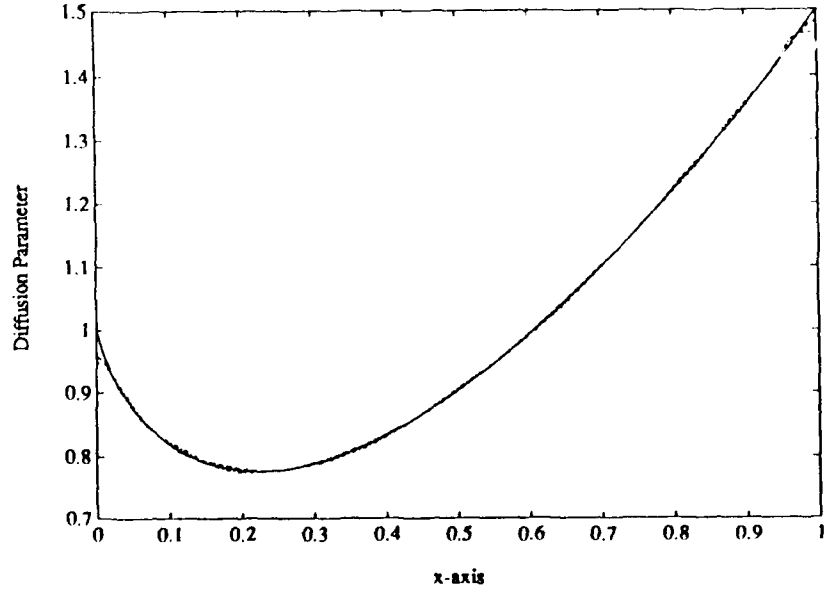

Figure 8. True and Approximate Diffusion Parameters for Example 6.3 with $\alpha = 10^{-8}$ $\cdot - \cdot$ $(M_x = 16)$, $- - -$ $(M_x = 24)$, —— (True).

Example 6.4 In this example, the parameter to be recovered is the shifted Gaussian function $p(x) = 1 + \frac{1}{4}e^{-40(x-\frac{1}{3})}$. When combined with the state solution, this dictates the choices $\alpha = \beta = \gamma = \frac{3}{2}$, and $\delta = 1$ for the sinc decay parameters as specified by (3.18). Pseudo-random noise is added to the data in the manner described in Example 6.2. As seen in Figure 9, the Tikhonov parameter values $\alpha = 10^{-5}$ through $\alpha = 10^{-8}$ yield points $(\|K(\vec{c}_\alpha) - \vec{d}\,\|, \|R\vec{c}_\alpha\|)$ in the "corner" of the L-curve. For $M_x = 16$, the uniform errors obtained with $\alpha = 10^{-3}$, $\alpha = 10^{-8}$, and $\alpha = 10^{-10}$ are reported in Table 4 with corresponding plots of the true and approximate parameters shown in Figure 10. As indicated by the numerical results, the "corner" value $\alpha = 10^{-8}$ provides a good choice for the regularization parameter whereas $\alpha = 10^{-3}$ causes too much smoothing. The error contributions due to the smaller singular values become quite apparent at $\alpha = 10^{-10}$ thus reiterating the L-curve observation that this Tikhonov value does not provide enough regularization or smoothing for the problem.

	$\alpha = 10^{-3}$	$\alpha = 10^{-8}$	$\alpha = 10^{-10}$
$\|p_U(\ell)\|$	$.7710 - 1$	$.4109 - 1$	$.6805 - 1$

Table 4. Errors on the Uniform Grid U with $M_x = 16$ in Example 6.4.

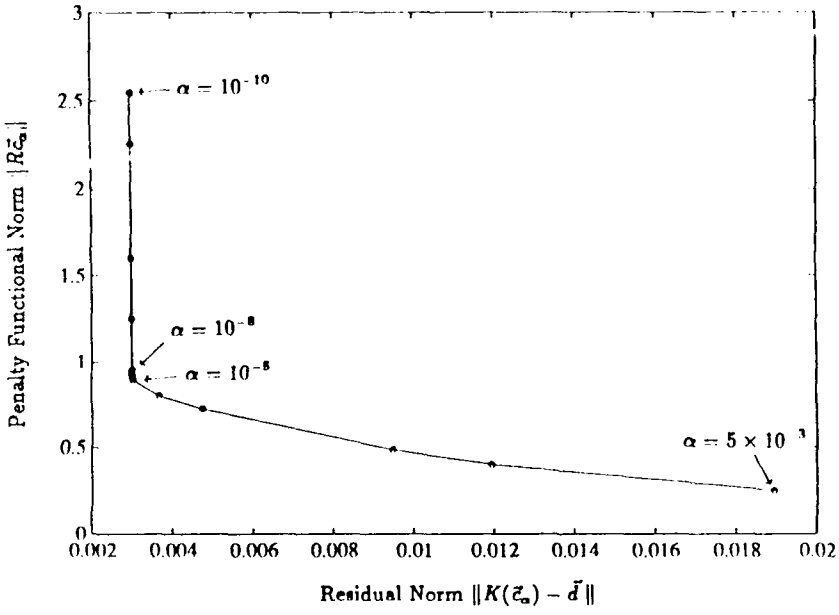

Figure 9. The Tikhonov L-curve for Example 6.4.

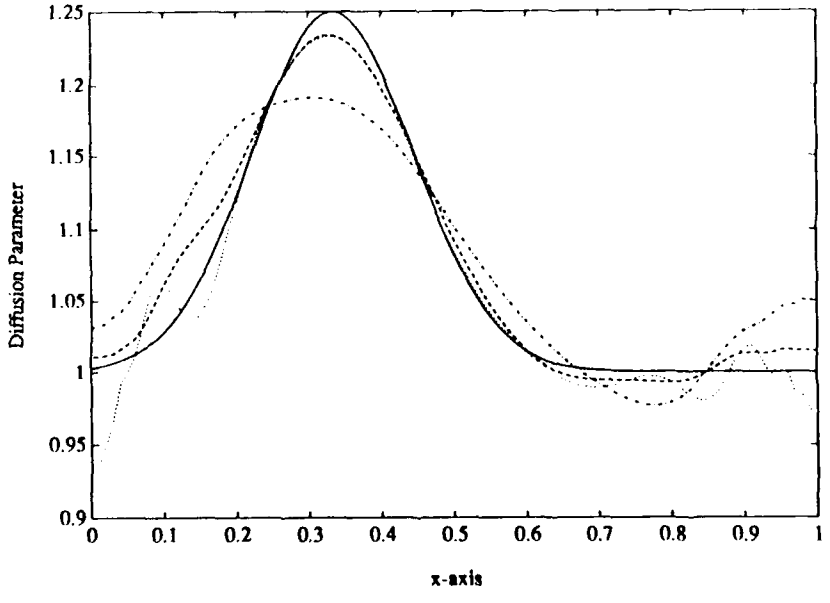

Figure 10. True and Approximate Diffusion Parameters for Example 6.4 with $M_x = 16$ $- \cdot -$ $(\alpha = 10^{-3})$, $- - -$ $(\alpha = 10^{-8})$, $\cdots$ $(\alpha = 10^{-10})$, —— (True).

References

[1] Banks H T and Kunisch K 1989 *Estimation Techniques for Distributed Parameter Systems* (Boston: Birkhäuser)

[2] Dennis J E and Schnabel R B 1983 *Numerical Methods for Unconstrained Optimization and Nonlinear Equations* (Englewood Cliffs, NJ: Prentice Hall)

[3] Doyle R R 1990 Extensions to the development of the Sinc-Galerkin method for parabolic problems *PhD Dissertation, Montana State University*

[4] Gill P E, Murray W and Wright M H 1981 *Practical Optimization* (New York: Academic Press)

[5] Golub G H, Nash S and VanLoan C 1979 A Hessenberg-Schur method for the problem $AX + XB = C$ *IEEE Trans. Automat. Control* 24 909-913

[6] Golub G H and VanLoan C 1989 *Matrix Computations*, 2nd ed. (Baltimore: Johns Hopkins University Press)

[7] Hansen P C 1990 Analysis of discrete ill-posed problems by means of the L-curve, submitted to *SIAM Rev.*

[8] Jonca K 1988 Numerical solution of a nonlinear Fredholm integral equation of the first kind *PhD Dissertation, Montana State University*

[9] Jonca K and Vogel C R 1989 Numerical solution to the magnetic relief problem, in *Transport Theory, Invariant Imbedding, and Integral Equations (Lecture Notes in Pure and Applied Mathematics)* ed P Nelson *et al* (New York: Marcel Dekker) 379-391

[10] Lewis D L, Lund J and Bowers K L 1987 The space-time Sinc-Galerkin method for parabolic problems *Int. J. Numer. Methods Eng.* 24 1629-1644

[11] Lund J 1986 Symmetrization of the Sinc-Galerkin method for boundary value problems *Math. Comp.* 47 571-588

[12] Lund J and Vogel C R 1990 A fully-Galerkin method for the solution of inverse problems in parabolic partial differential equations *Inv. Prob.* 6 205-217

[13] McArthur K M, Smith R C, Bowers K L and Lund J The Sinc-Galerkin method for parameter dependent self-adjoint problems, to be submitted to *J. of Comp. and Appl. Math.*

[14] Moler C B and Stewart G W 1973 An algorithm for generalized matrix eigenvalue problems *SIAM J. Numer. Anal.* 10 241-256

[15] Seidman T I and Vogel C R 1989 Well-posedness and convergence of some regularization methods for nonlinear ill-posed problems *Inv. Prob.* 5 227-238

[16] Stenger F 1976 Approximations via Whittaker's cardinal function *J. Approx. Theory* 17 222-240

[17] Stenger F 1979 A Sinc-Galerkin method of solution of boundary value problems *Math. Comp.* 33 85-109

[18] Stenger F 1981 Numerical methods based on Whittaker cardinal, or sinc functions *SIAM Rev.* 23 165-224

[19] Tikhonov A N and Arsenin V Y 1977 *Solutions of Ill-Posed Problems* (New York: Wiley)

[20] Wahba G 1977 Practical approximate solutions to linear operator equations when the data are noisy *SIAM J. Numer. Anal.* 14 651-667

Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
NASA CR-187567 ICASE Report No. 91-43		

4. Title and Subtitle	5. Report Date
SINC-GALERKIN ESTIMATION OF DIFFUSIVITY IN PARABOLIC PROBLEMS	May 1991
	6. Performing Organization Code

7. Author(s)	8. Performing Organization Report No.
Ralph C. Smith Kenneth L. Bowers	91-43
	10. Work Unit No.
	505-90-52-01

9. Performing Organization Name and Address	11. Contract or Grant No.
Institute for Computer Applications in Science and Engineering Mail Stop 132C, NASA Langley Research Center Hampton, VA 23665-5225	NAS1-18605
	13. Type of Report and Period Covered
12. Sponsoring Agency Name and Address	Contractor Report
National Aeronautics and Space Administration Langley Research Center Hampton, VA 23665-5225	14. Sponsoring Agency Code

15. Supplementary Notes

Langley Technical Monitor:
Michael F. Card

Submitted to Inverse Problems

Final Report

16. Abstract

A fully Sinc-Galerkin method for the numerical recovery of spatially varying diffusion coefficients in linear partial differential equations is presented. Because the parameter recovery problems are inherently ill-posed, an output error criterion in conjunction with Tikhonov regularization is used to formulate them as infinite-dimensional minimization problems. The forward problems are discretized with a sinc basis in both the spatial and temporal domains thus yielding an approximate solution which displays an exponential convergence rate and is valid on the infinite time interval. The minimization problems are then solved via a quasi-Newton/trust region algorithm. The L-curve technique for determining an appropriate value of the regularization parameter is briefly discussed, and numerical examples are given which demonstrate the applicability of the method both for problems with noise-free data as well as for those whose data contains white noise.

17. Key Words (Suggested by Author(s))	18. Distribution Statement
Sinc-Galerkin Method, Numerical estimation of diffusion coefficients	64 - Numerical Analysis Unclassified - Unlimited

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of pages	22. Price
Unclassified	Unclassified	36	A03